Gynecologic Oncology 132 (2014) 343–350

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

A distinctive ovarian cancer molecular subgroup characterized by poor prognosis and somatic focal copy number amplifications at chromosome 19

Chang Ohk Sung [a,*], In-Hye Song [a], Insuk Sohn [b,**]

[a] Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[b] Samsung Cancer Research Institute, Seoul, Republic of Korea

HIGHLIGHTS

- Consensus clustering based on the focal copy number alteration reveals a distinctive molecular subgroup.
- The subgroup characterized by amplification at chromosome 19 is independently associated with poor prognosis.
- The subgroup shows significant tendency toward mutual exclusivity with patients with *BRCA1/2* mutations.

ARTICLE INFO

Article history:
Received 28 August 2013
Accepted 30 November 2013
Available online 8 December 2013

Keywords:
Ovary
Copy number
Serous
Carcinoma
Prognosis
Clustering

ABSTRACT

Objective. High-grade serous ovarian cancer (HGS-OvCa), the most common epithelial ovarian cancer, is very complex and heterogeneous at the molecular level. The identification of intrinsic HGS-OvCa subgroups characterized by specific molecular alterations and aggressive behavior could improve patient treatment.

Methods. High-resolution copy number data for 560 HGS-OvCa patients and gene expression data obtained from the TCGA database were analyzed to identify distinct molecular subgroups based on significant focal somatic copy number alterations (SCNAs).

Results. Using unsupervised consensus clustering, a subgroup accounting for 26.8% of the patients (150/560 patients) characterized by focal somatic copy number amplification at chromosome 19 was identified. The subgroup was independently associated by multivariate Cox regression analysis with poor overall (HR, 1.61; $P = 0.001$) and progression-free survival (HR, 1.36; $P = 0.036$). The specific focal SCNA locations were 19p13.2, 19p13.12, 19p13.11, 19q12, 19q13.12, and 19q13.2. The differential gene expression signature of the subgroup compared with that of the remaining patients also suggested that chromosome 19 was the mainly amplified region. The clinical significances of subgroup 2 were validated in independent data sets using the gene expression signature characteristics. In addition, the subgroup had a tendency toward mutual exclusivity with patients with *BRCA1/2* mutations. The most significantly altered pathway of the subgroup was the cyclin and cell cycle regulation pathway.

Conclusion. A unique molecular subgroup associated with poor survival was identified based on focal SCNAs and could aid the further molecular classification of ovarian cancers.

Introduction

Ovarian cancers are histologically classified into four major subtypes: serous, clear cell, endometrioid, and mucinous. Of these, serous carcinoma is the most common epithelial ovarian cancer, accounting for almost two-thirds of ovarian cancer deaths [1]. Serous ovarian cancers are further divided into two distinct subtypes: low-grade (type I) and high-grade (type II) serous carcinomas. High-grade serous ovarian cancer (HGS-OvCa) is much more common than its low-grade counterpart and shows more aggressive behavior.

The current standard treatment for HGS-OvCa involves aggressive tumor cytoreductive surgery followed by platinum-based multiagent chemotherapy. HGS-OvCa is usually platinum sensitive but approximately 30% of affected patients exhibit platinum resistance and a more aggressive disease progression [2,3], and it is difficult to predict which patients will show a poor prognosis after standard treatment. These heterogeneities suggest the presence of intrinsic HGS-OvCa

* Correspondence to: C.O. Sung, Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro, 43-gil, Songpa-Gu, Seoul 138-736, Republic of Korea. Fax: +82 2 472 7898.
** Correspondence to: I. Sohn, Samsung Cancer Research Institute, 50 Ilwon-dong, Gangnam-gu, Seoul 135–710, Republic of Korea. Fax: +82 2 2148 9819.
E-mail addresses: co.sung@amc.seoul.kr (C.O. Sung), insuks@gmail.com (I. Sohn).

0090-8258/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ygyno.2013.11.036

molecular subtypes. Hence, the identification of the intrinsic HGS-OvCa subgroup characterized by a specific molecular alteration associated with more aggressive behavior could lead to a more successful treatment outcome. Currently, many studies of HGS-OvCa are focused at the single gene level, and little is known about the molecular classification of this disease. Resolving this important issue and increasing the number of trials of personalized cancer therapies for HGS-OvCa would likely improve patient outcomes.

A number of chromosomal aberrations and individual gene alterations have been reported in ovarian cancer including the association of a BRCA gene mutation with an improved chemotherapy response and survival outcome [4–6]. Hence, patients with BRCA gene mutation are an important ovarian cancer subgroup. However, a considerable number of HGS-OvCa patients have no germline or somatic BRCA gene mutations. In addition, other somatic mutations, with the exception of *TP53*, occur very rarely in this cancer [7]. In contrast, HGS-OvCa shows apparent DNA copy number changes, either germline or sporadic [8]. Therefore, because changes in copy number are a distinctive and consistent feature of HGS-OvCa, an altered copy number is emphasized more than mutational changes when establishing treatment strategies for this disease [8,9].

Copy number alterations in a genome cause the cell to have an abnormal number of one or more DNA regions. Copy number alterations can be divided into two categories: focal events and broad events. A focal event is an alteration in only a small proportion of the chromosome arm. These focal chromosomal aberrations enable easier identification of the driver genetic alterations than broad chromosomal aberrations and focal somatic copy number alterations (SCNAs) that entail an acquired increase in the copy number of a restricted region of the genome and play central roles in oncogenesis [10,11].

In our present study, we obtained and analyzed high-resolution copy number data for 560 HGS-OvCa patients from The Cancer Genome Atlas (TCGA) to identify distinct molecular subgroups based on significant focal SCNAs.

Materials and methods

Copy number data and analysis

Segmented copy number data generated with an Affymetrix Single Nucleotide Polymorphism (SNP) 6.0 array were obtained from the TCGA database (http://tcga-data.nci.nih.gov) for 560 patients with HGS-OvCa. The segmented copy number data included the somatic copy number but not the germline copy number. These data were built on an hg19 reference genome. The procedures, normalization, and processing methods for the data have been described previously [7]. Clinical information including age at diagnosis, overall survival (OS), progression-free survival (PFS), platinum response (sensitive or resistant), tumor grade, tumor stage, and *BRCA1* and *BRCA2* mutation status were also obtained. The *BRCA1/2* mutation status included both somatic and germline mutations. The parameters used have been defined in detail in previous reports [7,12,13]. All patients had advanced (FIGO stage $\geq$ 2) HGS-OvCa. The downloaded segmented copy number data were analyzed with GISTIC2 [11,14] to identify significant focal SCNAs. Thresholds for significant focal copy number alterations were as follows: amplification and deletion threshold, 0.1; cap values, 1.5; broad length cutoff, 0.7; confidence level, 0.95; joint segment size, 4; level peel-off, 1; and maximum sample segments, 2000. Details of each parameter have been described previously [11].

Gene expression and mutation count data

Normalized and unified gene expression data from expression microarrays for 489 patients were obtained from the TCGA database. The procedures, normalization, and processing methods have been detailed previously [7]. The BRBarray tool [15] was used to obtain the differentially expressed genes between subgroups that were identified on consensus clustering with $P < 0.0001$ and FDR < 0.05 thresholds. Gene ontology and pathway analysis with KEGG and BioCarta datasets were performed using the BRB-Array and DAVID bioinformatic tools [15–17]. The mutation count for each sample was calculated in somatic mutation data from whole exome sequencing of 311 patients downloaded from the TCGA data portal.

Consensus clustering with significant focal SCNA data

A consensus hierarchical, k-means, non-negative matrix factorization (NMF) clustering with iterative feature selection was performed. Consensus clustering is a resampling-based procedure that repeatedly samples a subset of the samples and then uses clustering to find intrinsic groupings [18,19]. Consensus clustering records the proportion of resamplings in which pairs of tumors were in the same clusters [18,19]. NMF, an algorithm based on decomposition by parts that can reduce the dimension of the data, is also an efficient method for the identification of distinct molecular patterns and provides a powerful method for class discovery [20]. These algorithms have been described in previous reports [7,20]. Consensus clustering analyses were performed using the GenePattern software of the Broad Institute with the "ConsensusClustering" and "NMFConsensus" modules and pipelines [21].

Statistical analysis

A *t*-test was used to evaluate the differences in the means for continuous variables between the two groups. A chi-square test or Fisher's exact test was used to test associations between two categorical groups. The OS and PFS values were determined using the Kaplan–Meier method, and survival curves were compared with the log-rank test. The Cox proportional hazards model was used to evaluate the relationships between the subgroups and the clinicopathological factors affecting the OS and PFS. Hazard ratio (HR), along with 95% confidence intervals (CIs), was assessed for each factor. The odds ratio was calculated with the Fisher's exact test to indicate the likelihood of mutually exclusivity or co-occurrence between two events [22]. All tests were two-sided and *P* values less than 0.05 were considered statistically significant. Statistical analysis was performed with Stata/IC statistical software (version 12; StataCorp, TX) and the R program (version 2.14.2; www.r-project.org).

Results

HGS-OvCa subgroup identification based on significant focal SCNAs

By using the actual copy number data generated after GISTIC2.0 analysis, we performed consensus clustering with three different algorithms (hierarchical, k-means, and NMF) to identify intrinsic subgroups among the HGS-OvCa patient cohort. Consensus clustering analyses suggest the presence of two ($k = 2$) or three subgroups ($k = 3$) of tumors in HGS-OvCa (Fig. 1A). The consensus was prominently decreased in four or more subtypes. We then tested the clinical significance of the three subgroups identified by consensus clustering with the three different algorithms. We found that subgroup 2 showed a poorer OS and PFS in all three different consensus clustering methods—hierarchical (Fig. 1B), k-means (Fig. 1C), and NMF (Fig. 1D)—than the other subgroups.

HGS-OvCa subgroup 2 is characterized by focal amplifications at chromosome 19

Of the three HGS-OvCa subgroups identified by the three different consensus clustering methods, subgroup 2 was found to be highly concordant in all clustering methods (Fig. 2A). Patients in subgroup 2 had